

# 数据挖掘在教育信息化中的应用研究

高晓佳

(吉林建筑工程学院 城建学院, 吉林 长春 130111)

**摘要:**随着数据挖掘理论在各个领域中的应用日益广泛,利用数据挖掘理论挖掘教育信息中的有价值信息就越发的体现出其必要性。该文从挖掘主题、数据库设计等几方面对教育决策系统进行了研究和设计,提出许多极具价值的辅助性建议,从而优化了高校的教学管理系统。

**关键词:**数据挖掘;挖掘主题;数据库;教育决策

**中图分类号:**TP393 **文献标识码:**A **文章编号:**1009-3044(2012)05-1199-02

**Data Mining in Education in the Application of Information Technology**

GAO Xiao-jia

(Jilin Architecture and Civil Engineering Institute, Institute of Urban Construction, Changchun 130111, China)

**Abstract:** With the theory of data mining applications in various fields, increasingly widespread use of data mining theory education information mining the more valuable information in the embodiment of its necessity. In this paper, mining theme, database design and other aspects of educational decision-making system design and research, The recommendations put forward many valuable auxiliary in order to optimize the universities' teaching management system.

**Key words:** data mining; mining theme; data base; educational decision-making

## 1 数据挖掘技术在教育信息化中的作用

随着信息技术和计算机技术的不断发展,计算机已经成为人们日常管理数据的必备工具。各个高校也开始使用计算机来管理教育信息。信息技术与教学结合的日益紧密,各高校不再满足于单纯的利用计算机管理数据,而是希望透过大量数据挖掘出数据后面所隐藏的信息和规律,从而为高校的教学管理提出一些具有建设性的建议<sup>[1]</sup>通过对目前各大高校的教学模式的研究,数据挖掘对高等教育的作用主要可以总结为如下几点:

### 1) 教学大纲的优化

教学大纲体根据教学计划,规定了每个学生在大学学习期间必须掌握的理论知识、实际技能和基本技能,也规定了教学进度和教学方法的基本要求。通过数据挖掘技术可以发现各课程之间的联系,对教学大纲中的课程设置、教学计划等进行优化。

### 2) 教学改革

授课教师的教学方式、所使用的教学设备直接关系到学生的成绩。利用数据挖掘技术挖掘出教学方法与学生成绩之间的规律就可以为各高校的教学改革提供有力依据。

## 2 教育决策系统挖掘主题设计

通过对当前各高校教育信息的研究,现将决策系统中的主题设计成如下五个:课程关联、课程类别关联、学生基本信息关联、课程与基本信息关联和教学模式关联<sup>[2]</sup>:

### 2.1 课程关联

该项关联主要用于挖掘各科目之间的关系,以优化教学大纲中各科目之间的衔接。但根据对现行各大高校成绩管理的研究发现,不同类别的课程的成绩形式并不完全一致,如:考试课一般按百分制而考察课通常是按优、良划分等级。为此在进行该项挖掘之前需要对各科成绩进行统一化。

通过该主题挖掘所得的结果形式如下:

数字逻辑 = A  $\Rightarrow$  计算机组成原理 = A (0.3, 0.9)

### 2.2 课程类别关联

本主题主要用于挖掘各不同类别课程之间的关系,用以安排各不同类别课程的侧重点。目前各高校主要按两种不同的标准对课程分类:一种按课程内容分为:公共基础课、专业基础课、专业课、实验和实践课;第二种按课程性质分为必修课、任意选修课和限定选修课。

收稿日期:2012-01-22

作者简介:高晓佳(1981-),男,吉林长春人,吉林建筑工程学院城建学院,助教,硕士,研究方向为计算机软件与理论。

本栏目责任编辑:梁书

由于每一类科目中不止一科,为了能综合体现出该类课程的整体状况,我们采用加权求平均的方法。权值主要依据该门课程的学分。

最终的挖掘结果形式如下:专业课综合成绩 = B  $\Rightarrow$  实验课综合成绩 = B (0.6, 0.9)。

### 2.3 学生基本信息的关联挖掘

本主题中所涉及的内容主要是学生们的自然信息,如:姓名,性别,民族等。以及学生已经结业的各门科目的成绩。

这些基本信息记录方式比较多样化,尤其是自然信息,为此我们需要在对数据进行挖掘之前需要对这些数据进行统一的标准化处理<sup>[5]</sup>下面对几个比较主要属性的标准化方式进行定义:

1)学生综合成绩=公共基础课综合成绩\*G1%+专业基础课综合成绩\*G2%+专业课综合成绩\*G3+实践环节综合成绩\*G4%

G为各类课程的权,要求G1+G2+G3+G4=100,其具体的取值可以根据用户和所涉及的挖掘内容进行调整。

各类别的综合成绩的求得可参考学生综合成绩。

结果的形式为:80分以上为优,60~79为良,低于60为差。

2)实践能力=实验课综合成绩\*G1+课程设计综合成绩\*G2+社会实践综合成绩\*G3,其中G1、G2、G3的定义可参考学生综合成绩评定中的权值。结果的定义形式为:80分以上为优,60~79为良,低于60为差。

3)通过对各高校学生信息的统计发现在学生自然信息中父母的受教育程度对孩子的成绩影响相对比较明显,在该项中将以父母的学历来产生相应的挖掘规则。

学历等级:博士:100,硕士研究生:85,高等教育:60,其他0

父母综合教育程度=(父亲学历的分数+母亲学历的分数)/2

结果的定义形式为:80分以上为优,60~79为良,低于60为差。

本项中可产生的关联规则形如:

父母教育程度=优 $\wedge$ 实践能力=优  $\Rightarrow$  综合成绩=优 (0.4, 0.8)

### 2.4 课程与基本信息关联挖掘

以往的就业基本都是用人单位对学生进行相关职业技能的考核,而学校往往对这些信息了解的并不详细,使得在进行就业指导 and 推荐的时候流于形式。本项主题可以根据学生的成绩和用人单位感兴趣的内容挖掘出某一名同学是否能够胜任某一项工作。

挖掘结果形式为:

软件建模分析与设计=A $\wedge$ 面向对象程序设计=A  $\Rightarrow$  工作=软件 (0.4, 0.7)。

### 2.5 教学模式关联挖掘

为了提高教学效果,优化教育资源,通常期望能够得到教学方法、教学设施等于教学效果之间的联系。本主题所要挖掘的内容即为教学方法和教师信息与教学效果的联系。从而优化教学方法。

其最终挖掘结果形式如下:教师职称=教授  $\Rightarrow$  平均分=A (0.3, 0.7)。

## 3 系统数据库设计

### 3.1 原始数据库

该数据库用来存放于挖掘主题相对应的一些基本数据和属性。通过对如上五个挖掘主题的总结可以将本数据库设计为如78个表:

1)学生自然信息表:存放学生的自然信息,如:姓名,性别,出生日期,身份证号,所学专业等。

2)课程表:存放学生在校期间应结业的科目,包含:课程名称,课程编号,学期,学分等信息。

3)成绩表:存放学生各科原始成绩。

4)规范化标准表:该表中存放用来将原始成绩规范化的各种标准。如:考察课为优良中及。

5)教学方式表:该表中存放的为授课教师的信息和教学设备等教学辅助设施。如:教师的自然信息、授课方式为演示、采用多媒体教学等。

6)系统代码表:存放了不同字段的属性值的类型,值域以及安全性设置等信息。

7)用户权限表:存放了可使用该系统的不同级别用户的登录信息以及权限。

## 4 挖掘库设计<sup>[6]</sup>

由于原始信息记录形式多样化,在进行数据挖掘前需要对原始信息进行标准化处理,本挖掘库中存放的即为转换后的数据。

1)学生基本自然信息标准化表:存放经过标准化处理后的学生的自然信息。

2)教学方式标准化中间表:存放在标准化过程中所得到的教师、教学方法等信息。

3)教学模式挖掘表:存放标准化后的教学模式信息。

4)阈值表:存放用户所需要的本次挖掘信息最小支持度和置信度的值。

5)候选集表:存放每次搜索数据库时产生的符合最小支持度和置信度的集合。

6)课程主题频繁项集表:存放课程主题的频繁项集。

7)课程类别频繁项集表:存放课程类别的频繁项集。

8)学生基本信息主题频繁项集表:用以存放与学生基本信息相关主题的频繁项集。

(下转第1202页)

构的职业培训体系嵌入学历教育体系、将职业资格认证体系嵌入课程体系、将行业标准嵌入课程标准和教学大纲、将企业文化嵌入校内教育环境”等“四个嵌入”作为专业改革的重点,实施了校企共育、工学结合高技能人才培养模式。与此同时,还需要做好以下几点:

1) 建好实习基地,强化学生的实践教学,提高学生的动手能力。

信工程设计与施工专业的学生需要有较高综合分析和实践运用能力,我们将在现有的3个校外实习基地的基础上,与多家通信施工企业等合作的基础上增加3个校外实习基地,以后每年增加不少于1个。与各实习基地达成长期协作、关系融洽的合作关系。

2) 在教学过程中,要创设工作情境,强化实际操作训练;要紧密结合职业技能证书的考核,例如施工必须的登高证、电工证等。

3) 在教学过程中,尽可能采用多媒体教学和实物教学,通过实际演练理解理论知识和强化技能;同时尽量采用小班上课。

4) 加强对外交流,为学生创造施展才华的机会。

坚持开放式办学,坚持教育社会化,为此我们将积极创造条件,加强国际交流与合作,不断开辟国际市场,探索国际联合办学的新途径。增强学生与其它国家学生的交流与合作,培养学生的综合素质;聘请社会和知名院校专家、教授举办讲座等,为学生创造一个良好的学习平台,为今后继续深造创造条件。

#### 4 结束语

课程建设的内容十分丰富,远不止课程设置和人才培养模式这两个方面,我们必须从市场需求、课程特点和高职院校教学规律出发,科学、合理地建设课程,才能更好地培养通信工程设计与施工人才、才能有效地实施课程,也才能把课程建设搞好。

---

(上接第1200页)

9)课程基本信息主题频繁项集表:用于存放各门课程与基本信息相关主题的频繁项集。

10)教学模式主题频繁项集:用以存放教学模式主题的频繁项集。

#### 5 总结

本文通过对各个高校的教学管理模式的研究发现,利用数据挖掘技术对当前教学信息中的一些潜在的规则进行挖掘具有极大的价值,不仅可以优化当前高校教学管理并且可以为招生就业,教学改革等提出许多富有建设性的辅助性建议。

#### 参考文献:

- [1] 魏萍萍,王翠茹.数据挖掘技术及其在高校教学系统中的应用[J].计算机工程.2004,29(11):87-89.
- [2] 李红艳.数据挖掘及其运用于教学评价的设想[J].襄樊职业技术学院学报,2005,2(1):42-45.
- [3] 李向东,宋一中.试卷分析指标及其计算方法研究[J].中国高等教育评估,2004(2):56-58.
- [4] 李福华.高等学校办学规模与办学条件分析[J].中国高等教育评估,2003(3):27-30.
- [5] 姜继红,王毅杰.地方性高校毕业生就业意向的调查分析[J].教育发展研究,2004 (5):83-86.
- [6] 曲春锦/Apriori-TIDS算法设计及其在教育决策信息挖掘中的应用[D].上海:上海海事大学,2005.